





[illegible]

```

/clone="IMAGE:5424766"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_43"
/notes="Organ: eye; Vector: pOR7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dt priming, _directionally_
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAC(G) . Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"

```

Query Match	75.1%	Score 743:	DB 12:	Length 814:
Best Local Similarity	98.6%	Pred. No. 2,96-82:		
Matches 802:	Conservative 0:	Mismatches 5:	Indels 6:	Gaps 1
QY 88	CCGAGCCGAGTGCCTCCGCTCTCCGTGCCCCGCGCTGCGCTGAGCTGAGCTCAGCCGACCTGG	14		
Db 2	CCGAGCCGAGTGCCTCCGCTCTCCGTGCGCCCGCGCTGCGCTGAGCTGAGCTCAGCCGACCTGG	61		
QY 148	GCGCCGCTTCGCGACATGAGCTCTTCTCTGAGGAGAGCGGTGCCCCCAATGCGAGCGGCG	20		
Db 62	GCGCCGCTTCGCGACATGAGCTCTTCTCTGAGGAGAGCGGTGCCCCCAATGCGAGCGGCG	122		
QY 208	GAGAGCAGTGCCTGCTTCCCGCCGAGGACAGCGCGCTGTGAGCAATCTTCTGAGCGGCT	26		
Db 122	GAGAGCAGTGCCTGCTTCCCGCCGAGGACAGCGCGCTGTGAGCAATCTTCTGAGCGGCT	18		
QY 268	CCATGCGGAGACACCCGCGCTGCGAAGCTTGAAGCTTGACCTTGAGACGCGGACG	32		
Db 182	CCATGCGGAGACACCCGCGCTGCGAAGCTTGAAGCTTGACCTTGAGACGCGGACG	241		
QY 328	GCGATTCTAATGATGACCTGCGAGACAGGCCACGCTTTGAGCCACCTTGCGCGCTCATCC	38		
Db 242	GCGATTCTAATGATGACCTGCGAGACAGGCCACGCTTTGAGCCACCTTGCGCGCTCATCC	303		
QY 388	AGGCAAGAGAGCGCTGACCTTGAGCACTTCACTGCGGCTTACTCCGCGCTGAGCCTGAGCC	44		
Db 302	AGGCAAGAGAGCGCTGACCTTGAGCACTTCACTGCGGCTTACTCCGCGCTGAGCCTGAGCC	362		
QY 448	TGGCGGTGCGCGCGGACCGGCGCGGTGTGATCTTGCGAGGTGAGCGCGACGCCCGGAGC	507		
Db 362	TGGCGGTGCGCGCGGACCGGCGCGGTGTGATCTTGCGAGGTGAGCGCGACGCCCGGAGC	421		
QY 508	TGGGACGAGCCCTGTGTGAGAGGACGCGGACGAGACAAAGTGCACCTCCCGCTGAAGC	567		
Db 422	TGGGACGAGCCCTGTGTGAGAGGACGCGGACGAGACAAAGTGCACCTCCCGCTGAAGC	481		
QY 568	CCGCTTTGAGAGACCTTGACGAGCTGTGAGCGGCGGAGC--GAGGCCGAGACCTTTCAGCGT	625		
Db 482	CCGCTTTGAGAGACCTTGACGAGCTGTGAGCGGCGGAGC--GAGGCCGAGACCTTTCAGCGT	541		
QY 626	GAGCGCTGTGAGTGTGCGGACAAAGAGAACTGCTCCGCTCATCTGAGAGCGTGTGCTGAGC-	684		
Db 542	GAGCGCTGTGAGTGTGCGGACAAAGAGAACTGCTCCGCTCATCTGAGAGCGTGTGCTGAGC-	601		
QY 686	TGCTTGCGACCCCGAGGACATCTCTGCGCGTCTTCAGAGTCTCTGTGCGCGCGGAAAGTGTCTGC	744		
Db 602	TGCTTGCGACCCCGAGGACATCTCTGCGCGTCTTCAGAGTCTCTGTGCGCGCGGAAAGTGTCTGC	661		
QY 745	AACCTTCGAAAGGGAGCGTGTGCGCGCGAGTG-TGTGCGAAACCTTAAAGAAC-GCATCCG	802		
Db 662	AACCTTCGAAAGGGAGCGTGTGCGCGCGAGTGTGTGCGAAACCTTAAAGAACGCGCATCCG	721		
QY 803	GCGGACGTACAGGTCTACATACGCTCTGCGCGCTGTGAGGATGAGTACACTTGGCGCTT	862		
Db 722	GCGGACGTACAGGTCTACATACGCTCTGCGCGCTGTGAGGATGAGTACACTTGGCGCTT	781		
QY 863	CAAGATCTAGGCGCTGCGCCCTTAATGTAGTGGGCT	895		
Db 782	CAAGATCTAGGCGCTGCG--CCCTAATGTAGTGGGCT	813		

RESULT 4	CA777547/c	756 bp	mRNA	linear	EST 03-DEC-2004
LOCUS	CA777547/c				
DEFINITION	ip21g07.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6217933 3'				
	similar to TR:085769 085769 HYPOTHETICAL 24.8 KD PROTEIN. ;, mRNA				
sequence.					
ACCESSION	CA777547				
VERSION	CA777547.1	GI:26015422			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	1 (bases 1 to 756)				
	Melton, D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K., Lemstra, I., Searce, M., Bressell, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Page, D., Wylie, T., Martin, J., Blistrain, A., Schmitt, A., Theising, B., Ralder, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tagarelsky, R., Williams, T., Jackson, Y., and Bowers, Y.				
TITLE	Endocrine Pancreas Consortium				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue Endocrine Pancreas Consortium Harvard University, Howard Hughes Medical Institute Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138 Tel: 617-495-1812 Fax: 617-495-8557 Email: dmelton@biochem.harvard.edu Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@um.wustl.edu) Seq primer: -400p from Gibco High quality sequence stop: 462. Location/Qualifiers 1..756 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6217933" /tissue_type="Purified pancreatic islet" /lab_host="DH10B" /clone_id="HR85 islet" /note="Organ: Pancreas; Vector: plasmid SK(-); Site_1: Not; Site_2: XhoI; cDNA made by oligo-dt priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permut Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."				

FEATURES	source
Db	402 CTGGACCTGGGACACCTTCAACGGGCTACCTCCGGCTCGGACCTCGGACCTGGCCGGG 461
Qy	402 CTGGACCTGGGACACCTTCAACGGGCTACCTCCGGCTCGGACCTCGGACCTGGCCGGG 461
Db	576 CTGACACTGGGACCTTCAACGGGCTACTCCGGCTTGGCTTGGCTTGGCTTGGCTTGG 517
Qy	576 CTGACACTGGGACCTTCAACGGGCTACTCCGGCTTGGCTTGGCTTGGCTTGGCTTGG 517
Db	462 GACGGGCGGCTGGTGAACCTGCGAGGTGAGCGGACGCCGCCGGACCTGGACCGGCCCTTG 521
Qy	462 GACGGGCGGCTGGTGAACCTGCGAGGTGAGCGGACGCCGCCGGACCTGGACCGGCCCTTG 521
Db	516 GACGGGCGGCTGGTGAACCTGCGAGGTGAGCGGACGCCGCCGGACCTGGACCGGCCCTTG 457
Qy	516 GACGGGCGGCTGGTGAACCTGCGAGGTGAGCGGACGCCGCCGGACCTGGACCGGCCCTTG 457
Db	522 TGGAGGCGAGGCCGAGGCCGAGCACAAAGTCCAGCTTCCGGCTGGAAGCCGGCTTTGGAGAC 581
Qy	522 TGGAGGCGAGGCCGAGGCCGAGCACAAAGTCCAGCTTCCGGCTGGAAGCCGGCTTTGGAGAC 581
Db	456 TGGAGGCGAGGCCGAGGCCGAGCACAAAGTCCAGCTTCCGGCTGGAAGCCGGCTTTGGAGAC 397
Qy	456 TGGAGGCGAGGCCGAGGCCGAGCACAAAGTCCAGCTTCCGGCTGGAAGCCGGCTTTGGAGAC 397
Db	582 CTGGACGAGCTGCTGGGCGGCGGGGCGGGACCTTCCAGACGTGGCCGTGGTGGATGCG 641
Qy	582 CTGGACGAGCTGCTGGGCGGCGGGGCGGGACCTTCCAGACGTGGCCGTGGTGGATGCG 641
Db	396 CTGGACGAGCTGCTGGGCGGCGGGGCGGGACCTTCCAGACGTGGCCGTGGTGGATGCG 337
Qy	396 CTGGACGAGCTGCTGGGCGGCGGGGCGGGACCTTCCAGACGTGGCCGTGGTGGATGCG 337
Db	642 GACAAAGAGAACTGCTCCGCCCTACTACAGAGCGCTGCTGCACTGCTGCACTGCACTGCAAGGCG 701
Qy	642 GACAAAGAGAACTGCTCCGCCCTACTACAGAGCGCTGCTGCACTGCTGCACTGCAAGGCG 701
Db	336 GACAAAGAGAACTGCTCCGCCCTACTACAGAGCGCTGCTGCACTGCTGCACTGCAAGGCG 277
Qy	336 GACAAAGAGAACTGCTCCGCCCTACTACAGAGCGCTGCTGCACTGCTGCACTGCAAGGCG 277
Db	702 ATCTCTCGCCGCTCTCAAGAGTCTTGGCGGGGAAAGTGCTGGCACTTCCGAAAGGAGAC 761
Qy	702 ATCTCTCGCCGCTCTCAAGAGTCTTGGCGGGGAAAGTGCTGGCACTTCCGAAAGGAGAC 761
Db	276 ATCTCTCGCCGCTCTCAAGAGTCTTGGCGGGGAAAGTGCTGGCACTTCCGAAAGGAGAC 217
Qy	276 ATCTCTCGCCGCTCTCAAGAGTCTTGGCGGGGAAAGTGCTGGCACTTCCGAAAGGAGAC 217
Db	762 GTGGCGGGCGGAGTGTGGCGGAACTTAAAGCAAGCATCCGGCGGGGAGCTCAAGGGTCTAC 821
Qy	762 GTGGCGGGCGGAGTGTGGCGGAACTTAAAGCAAGCATCCGGCGGGGAGCTCAAGGGTCTAC 821
Db	216 GTGGCGGGCGGAGTGTGGCGGAACTTAAAGCAAGCATCCGGCGGGGAGCTCAAGGGTCTAC 157
Qy	216 GTGGCGGGCGGAGTGTGGCGGAACTTAAAGCAAGCATCCGGCGGGGAGCTCAAGGGTCTAC 157
Db	822 ATCAGCCTCTGCGCCCTCTGGGCGATGGACTCACTTGGCCTTCAAGATCTAGAGGCTGGCC 881
Qy	822 ATCAGCCTCTGCGCCCTCTGGGCGATGGACTCACTTGGCCTTCAAGATCTAGAGGCTGGCC 881
Db	156 ATCAGCCTCTGCGCCCTCTGGGCGATGGACTCACTTGGCCTTCAAGATCTAGAGGCTGGCC 97
Qy	156 ATCAGCCTCTGCGCCCTCTGGGCGATGGACTCACTTGGCCTTCAAGATCTAGAGGCTGGCC 97
Db	882 -CTAGTAGAGGGGTCCAGAGGAGGGTGGCTCTGGGAACCCGAGGAATTGACCTTGAGTTT 940
Qy	882 -CTAGTAGAGGGGTCCAGAGGAGGGTGGCTCTGGGAACCCGAGGAATTGACCTTGAGTTT 940
Db	96 GATGATGATGAGGTCTCAAGAGGAGGTGGCTCTGGGAACCCGAGGAATTGACCTTGAGTTT 37
Qy	96 GATGATGATGAGGTCTCAAGAGGAGGTGGCTCTGGGAACCCGAGGAATTGACCTTGAGTTT 37
Db	941 AAATTGAAAAATTAAGTGCGGCTGCGACACAAAAA 976
Qy	941 AAATTGAAAAATTAAGTGCGGCTGCGACACAAAAA 976
Db	36 AAATTGAAAAATTAAGTGCGGCTGCGACACAAAAA 1
Qy	36 AAATTGAAAAATTAAGTGCGGCTGCGACACAAAAA 1
RESULT 5	
LOCUS	BF663323 1006 bp mRNA linear EST 21-DEC-2000
DEFINITION	60214446.F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4297911 5',
ACCESSION	BF663323 mRNA sequence.
VERSION	BF663323
KEYWORDS	BF663323.1 GI:11937218
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	1 (bases 1 to 1006)
COMMENT	NIH-MGC http://mgs.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Louis M. Staudt, M.D., Ph.D. CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LICM152 row: m column: 16 High quality sequence stop: 745. Location/Qualifiers 1..1006 /organism="Homo sapiens" /mol_type="mRNA"

/db xref="taxon:9606"  
 /clone="IMAGE:4297911"  
 /tissue\_type="Primary B-cells from tonsils (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH\_MGC\_48"  
 /note="Organ: B-cells; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCAAGAG(9). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by Ling  
 Hong in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC library."

## ORIGIN

Query Match 72.9%; Score 721; DB 10; Length 1006;  
 Best Local Similarity 90.9%; Pred. No. 1.4e-79;

Matches 835; Conservative 0; Mismatches 75; Indels 9; Gaps 6;

```

QY 79 GGGCCATGACCCAGCGGTGCGCCCGCTCTCCGTGCGCGCCGCTGAGCCCTGAGCTGAG 138
Db 2 GGGCCATGACCCAGCGGTGCGCCCGCTCTCCGTGCGCGCCGCTGAGCCCTGAGCTGAG 61
QY 139 CCGCACTGGGCGCGCGCTTCCGCACTGCGCTTCTTCTGGGAGGCGGTGCGCCCATGAGC 198
Db 62 CCGCACTGGGCGCGCGCTTCCGCACTGCGCTTCTTCTGGGAGGCGGTGCGCCCATGAGC 121
QY 199 GAGGCGGCGAGAGAGTGTCTTCCCGCCGAGAGAGCGCGCTGTGGAGATCTTCTC 258
Db 122 GAGGCGGCGAGAGAGTGTCTTCCCGCCGAGAGAGCGCGCTGTGGAGATCTTCTC 181
QY 259 TGAAGCGCTTCATGCGGAGAGACCCGCGCGCTGCGAAGCTTGAAGCTTGAACCTTGAGC 318
Db 182 TGAAGCGCTTCATGCGGAGAGACCCGCGCGCTGCGAAGCTTGAAGCTTGAACCTTGAGC 241
QY 319 AGCGCGAGGGGATTTCTATGATGACTTCTCGAGAGAGAGCCAGCTCTTGGCCAACTTGAGC 378
Db 242 AGCGCGAGGGGATTTCTATGATGACTTCTCGAGAGAGAGCCAGCTCTTGGCCAACTTGAGC 301
QY 379 GGGTCATCAGAGCCAGAGAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 438
Db 302 GGGTCATCAGAGCCAGAGAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
QY 439 CCTTGGCGCTTCCGCTGCGCGCGAGAGCGGCGCGGTGTGACCTTGAAGTGAAGCGAGC 498
Db 362 CCTTGGCGCTTCCGCTGCGCGCGAGAGCGGCGCGGTGTGACCTTGAAGTGAAGCGAGC 421
QY 499 CCGCGAGAGTGGAGAGCGCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 558
Db 422 CCGCGAGAGTGGAGAGCGCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
QY 559 GGGTGAAGCCCGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 618
Db 482 GGGTGAAGCCCGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541
QY 619 TCGAGCTGCGCGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 678
Db 542 TCGAGCTGCGCGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
QY 679 TCGAGCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 737
Db 602 TCGAGCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661
QY 738 GTGTGCAACCTTCCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 796
Db 662 GTGTGCAACCTTCCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 721
QY 797 CATTCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 854
Db 722 CATTCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 781
QY 855 -TTGGCTTCAAG-ATCTAGAGGCTGGCCCTCTAGTAGTGGGCTCAGAGAGAGAGGTTG--C 909

```

```

Db 782 TTGAGCCTTCAAGGATTAAGAGGTGGCCCTTAAAGAGAGGTTCAGACGACGTTGGCC 841
QY 910 CTGGGAAACCCAGAGATTGACCTTGAATTAAATTCGAAATTAAGTGGGCTGGACA 969
Db 842 TCGGGAACCCCTAGAGATTGGCCCGCGTTAAATCCGACACTAGCTGTGGGACACACC 901
QY 970 CAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 988
Db 902 ATTAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 920

```

RESULT 6  
 BF664198  
 LOCUS  
 DEFINITION 602145812P1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4309298 5',  
 mRNA sequence.  
 BF664198  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNLI at:  
 http://image.llnl.gov  
 Plate: L1CM182 row: h column: 03  
 High quality sequence stop: 726.

## FEATURES

Location/Qualifiers  
 1..921

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4309298"  
 /tissue\_type="Primary B-cells from tonsils (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH\_MGC\_48"  
 /note="Organ: B-cells; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCAAGAG(9). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by Ling  
 Hong in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC library."

## ORIGIN

Query Match 72.4%; Score 716; DB 10; Length 921;  
 Best Local Similarity 97.4%; Pred. No. 5.9e-79;  
 Matches 781; Conservative 0; Mismatches 15; Indels 6; Gaps 5;

```

QY 79 GGGCCATGACCCAGCGGTGCGCCCGCTCTCCGTGCGCGCCGCTGAGCCCTGAGCTGAG 138
Db 2 GGGCCATGACCCAGCGGTGCGCCCGCTCTCCGTGCGCGCCGCTGAGCCCTGAGCTGAG 61
QY 139 CCGCACTGGGCGCGCGCTTCCGCACTGCGCTTCTTCTGGGAGGCGGTGCGCCCATGAGC 198
Db 62 CCGCACTGGGCGCGCGCTTCCGCACTGCGCTTCTTCTGGGAGGCGGTGCGCCCATGAGC 121
QY 199 GAGGCGGCGAGAGAGTGTCTTCCCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 258

```

```

Db      122 GAGGCGGCGAAGACAGTCTGCTTCCCGGAGAGACGCGCTGTGGAGATCTTC 181
QY      259 TGAGCGGCTCCATGCGGAGAGACCGGCGCTGCGAAGCTGAGGCTGACCTCGAGG 318
Db      182 TGAGCGGCTCCATGCGGAGAGACCGGCGCTGCGAAGCTGAGGCTGACCTCGAGG 241
QY      319 AGCCGAGGAGGATCTTATGATGACCTGCGAGCGAGCCCAAGCTTGGCCAACTGGG 378
Db      242 AGCCGAGGAGGATCTTATGATGACCTGCGAGCGAGCCCAAGCTTGGCCAACTGGG 301
QY      379 GAGCTATCCAGGCGCAAGAGGCGCTGAGACCTTCCAGGCTTACTCCGCTGG 438
Db      302 GAGCTATCCAGGCGCAAGAGGCGCTGAGACCTTCCAGGCTTACTCCGCTGG 361
QY      439 CCCTGAGCGCTGCGCTGCGCGCGAGCGGCGCTGAGTACCTGCGAGGAGGAGCGGCGAGC 498
Db      362 CCCTGAGCGCTGCGCTGCGCGCGAGCGGCGCTGAGTACCTGCGAGGAGGAGCGGCGAGC 420
QY      499 CCGCGAGAGTGGAGAGCGCCCTGTGTGAGAGCGAGCGGCGAGGAGCAAGATCGACTCC 558
Db      421 CCGCGAGAGTGGAGAGCGCCCTGTGTGAGAGCGAGCGGCGAGGAGCAAGATCGACTCC 480
QY      559 GAGCTAAGCGCGCTTGGAGAGCGGCGCTGAGAGCGGCGCTGAGTACCTGCGAGGAGGAGCGGCGAGC 618
Db      481 GAGCTAAGCGCGCTTGGAGAGCGGCGCTGAGAGCGGCGCTGAGTACCTGCGAGGAGGAGCGGCGAGC 540
QY      619 TCGAGCTGCGCGTGTGTGATGCGGAGCAAGAGAACTGCTCCGCTACTACGAGCGGCTGCC 678
Db      541 TCGAGCTGCGCGTGTGTGATGCGGAGCAAGAGAACTGCTCCGCTACTACGAGCGGCTGCC 599
QY      679 TCGAGCTGCTGCGAGCGGCGAGGAGCTTCTGCGCGCTGCTTCAAGTCTGTGGCGCGAGAGG 738
Db      600 TCGAGCTGCTGCGAGCGGCGAGGAGCTTCTGCGCGCTGCTTCAAGTCTGTGGCGCGAGAGG 659
QY      729 TCGTCAACCTCCGAAAGGAGAGCGTGGCGCGAGTGTGTGCGAAAGCTTAACGAGCGCA 798
Db      660 TCGTCAACCTCCGAAAGGAGAGCGTGGCGCGAGTGTGTGCGAAAGCTTAACGAGCGCA 718
QY      799 TCGCGCGGAGAGCTGAGGCTTACATCAAGCTCTGCGCGCTGCGAGGAGCTCACTGG 858
Db      719 TCGCGCGGAGAGCTGAGGCTTACATCAAGCTCTGCGCGCTGCGAGGAGCTCACTGG 775
QY      859 CTTCAAGATCTAGGCGCTGGCC 880
Db      776 GCTCAGAGATCTAGGCGCTGGCC 797

```

## RESULT 7

BM976177/c

LOCUS 714 bp mRNA linear EST 21-FEB-2003  
 DEFINITION UI-CF-EN1-acy-e-23-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone

ACCESSION BM976177  
 VERSION BM976177.1 GI:19593768

KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 714)  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene

TITLE Normalization and subtraction: two approaches to facilitate gene

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT 8889548

McCraty Lab  
 University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171  
 Email: paul-mccraty@uiowa.edu

## FEATURES

Location/Qualifiers  
 1..714  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-EN1-acy-e-23-0-UI"  
 /tissue\_type="Primary Lung Cystic Fibrosis Epithelial Cells"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-CF-EN1"  
 /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGGT.  
 TAG TISSUS=Human Lung Epithelial Cell Lines untreated LPS  
 6hr to LPS 24h  
 TAG LIB=UI-CF-EN1  
 TAG\_SEQ=CTGCTCAGGT"

## ORIGIN

Query Match 71.5%; Score 707.2; DB 12; Length 714;  
 Best Local Similarity 99.6%; Pred. No. 7.6e-79;  
 Matches 709; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY      278 GACCCGCGGCTGAGAGCTGAGGCTGAGACCTTGGAGACCGCGAGGAGATTCTAT 337
Db      714 GACCCGCGGCTGAGAGCTGAGGCTGAGACCTTGGAGACCGCGAGGAGATTCTAT 655
QY      338 GATGACCTGAGAGAGCGCCAGCTTGGAGCAAGCTTGGAGCGGCTCATCCAGGCGCAAGAA 397
Db      654 GATGACCTGAGAGAGCGCCAGCTTGGAGCAAGCTTGGAGCGGCTCATCCAGGCGCAAGAA 595
QY      398 GGGGCTGAGAGCTGGGAGCACTTCAAGGGCTACTCCGCTTGGCGCTTGGCGCTGCC 457
Db      594 GGGGCTGAGAGCTGGGAGCACTTCAAGGGCTACTCCGCTTGGCGCTTGGCGCTGCC 535
QY      458 CGGAGACGGGCGGTGGTGAAGCTGCGAGGTGAGAGCGGAGCGGAGGAGTGGAGCGGCC 517
Db      534 CGGAGACGGGCGGTGGTGAAGCTGCGAGGTGAGAGCGGAGCGGAGGAGTGGAGCGGCC 475
QY      518 CTTGTGAGAGCGAGGCGGAGCGAGAGCAAGATCGAAGCTTGGAGCGGCTGAGAGCGGCTTGA 577
Db      474 CTTGTGAGAGCGAGGCGGAGCGGAGCAAGATCGAAGCTTGGAGCGGCTGAGAGCGGCTTGA 415
QY      578 GACCTTGGAGAGAGCTTGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 637
Db      414 GACCTTGGAGAGAGCTTGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 355
QY      638 TCGGAGCAAGAGAGAACTGCTCCGCTACTAGAGCGGCTGCTGAGAGCGGAGCGGAG 697
Db      354 TCGGAGCAAGAGAGAACTGCTCCGCTACTAGAGCGGCTGCTGAGAGCGGAGCGGAG 295
QY      698 AGGCATCTTCCGCGCTCTCAAGAGTCTGTGGCGGAGAGGTGCTGCAACCTCCGAAAG 757

```

Db	294	AGGCATCTCTGCGCGTCTCAGAGTCTCTGTGGCGCGGAAAGTGCTCAACTCTCGAAAG	235
Qy	758	GGACGTGGCGCGCGAGTGTGTGGAAACCTTAAACGAACGATCCGGCGGGACGTCAAGGT	817
Db	234	GGACGTGGCGCGCGAGTGTGTGGAAACCTTAAACGAACGATCCGGCGGGACGTCAAGGT	175
Qy	818	CTACATCAACCTTCCTGCCCCCTGGGCGATGAGTCACTTGGCCTTCAAGATCTAAGGCTG	877
Db	174	CTACATCAACCTTCCTGCCCCCTGGGCGATGAGTCACTTGGCCTTCAAGATCTAAGGCTG	115
Qy	878	GCCCTAGTAGTGGGCTTCGAGGGAGGTTGCTGGGAACCCGAGAAATTGACCTGAGT	937
Db	114	GCCCTAGTAGTGGGCTTCGAGGGAGGTTGCTGGGAACCCGAGAAATTGACCTGAGT	55
Qy	938	TTTAATTTCGAAAAATTAAGTGGGGCTGGGACACAAAAAATTTTTTTTTTTTTTTT	989
Db	54	TTTAATTTCGAAAAATTAAGTGGGGCTGGGACACACAAAAAATTTTTTTTTTTTTTTT	3

RESULT 8	
CB852718	
LOCUS	
DEFINITION	CB852718 750 bp mRNA linear EST 22-APR-2001
ACCESSION	U1-CF-FN0-afl-p-08-0-U1.s1 U1-CF-FN0 Homo sapiens CDNA clone U1-CF-FN0-afl-p-08-0-U1 3', mRNA sequence. CB852718

VERSION	CB852718.1	GI:30047716
KEYWORDS	EST.	
SOURCE	Homo sapiens	
ORGANISM	Homo sapiens (human)	

REFERENCE  
AUTHORS  
TITLE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 750)  
Bonaldi,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene

discovery  
Genome Res. 6 (9), 791-806 (1996)  
MEDLINE  
97044477

PUBMED 8809548  
COMMENT  
Contact: McCray, PB

McCray Lab  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel.: 319 356 4866  
Fax: 319 356 7171  
Email: paul-mccray@uiowa.edu  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics ([www.resgen.com](http://www.resgen.com)) or from Open Biosystems  
([www.openbiosystems.com](http://www.openbiosystems.com)).  
Seq primer: M13 FORWARD  
POLYA=No.

```

FEATURES
source
location/Qualifiers
1..750
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FNO-at1-p-08-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_id="UI-CF-FNO"
/note="Organ: lung; Vector: pT73-Pac (pharmacia) with a
modified polylinker; Site 1: Ecor I; Site 2: Not I;
UI-CF-FNO is a subtracted cDNA library derived from two
normalized human lung epithelial cell libraries (EN1 and
DU1) The library was subtracted according to according to
1986. For additional information, contact:
bento-soares@uiowa.edu
TAG_SEQ=None found"

```

ORIGIN	
Query Match	70.8%; Score 699.8; DB 14; Length 750;
Best Local Similarity	97.3%; Pred. No. 6.1e-77;
Matches 732; Conservative	0; Mismatches 18; Indels 2; Gaps 2;

QY	148	GCGCGGCTTTGCGCAGTGGCTCTTCCCTGGGAGAGCGGTGGCCCCCAATGGCGAGCGGC	207
Db	1	GGGGCGCTCCAAACACCTCTCCGCCGACAGGAGAGCGGTGGCCCCCAATGGCGAGCGGC	60
QY	208	GAGAGCAGTGGCTTGTCTTCCCTCCGAGGACAGCGGCTGTGGCAGTATCTTGTAGCGGCT	267
Db	61	GAGAGCAGTGGCTTGTCTTCCCTCCGAGGACAGCGGCTGTGGCAGTATCTTGTAGCGGCT	120
QY	268	CCATGCGGAGGACACCGGAGGCTGGAGAGCTGAGAGGCTGACCCCTGAGAGCAGCGCAGG	327
Db	121	CCATGCGGAGGACACCGGAGGCTGGAGAGCTGAGAGGCTGACCCCTGAGAGCAGCGCAGG	180
QY	328	GGGATTTATATGATGACTGTGGACAGGCCAGCTCTTGGCCAACTGGCGGAGCTCATCC	387
Db	181	GGGATTTATATGATGACTGTGGAGAGGCCAGGCTTGGCCAACTGGCGGAGCTCATCC	240
QY	388	AGGCCAAGAGAGGCGCTGGACTTGGCACCCTTACAGGCTACTCCGCCCTTGACCCTTGACC	447
Db	241	AGGCCAAGAGAGGCGCTGGACTTGGCACCCTTACAGGCTACTCCGCCCTTGACCCTTGACC	300
QY	448	TGGCGCTGCGCGGAGCGGGCGGTGATCTGCGAGGTGGAGCGGCGACGCCCGGAGG	507
Db	301	TGGCGCTGCGCGGAGCGGGCGGTGATCTGCGAGGTGGAGCGGCGACGCCCGGAGG	360
QY	508	TGGGACGCGCCCTGTGAGGACGAGGCGGAGCGGAGGACAAAGATGACCTCCGAGCTGAAGC	567
Db	361	TGGGACGCGCCCTGTGAGGACGAGGCGGAGGCGGAGGACAAAGATGACCTCCGAGCTGAAGC	420
QY	568	CGGCTTTGAGAGCCCTTGGACGAGCTGTGGCGGCGGAGCGCGGACCTTGGACGTGG	627
Db	421	CGGCTTTGAGAGCCCTTGGACGAGCTGTGGCGGCGGAGCGCGGACCTTGGACGTGG	480
QY	628	CGGTGTGTATGCGGACAGAGAGAACTGTCCCGCTACTACAGAGCGTGTGCTGAGCTGC	687
Db	481	CGGTGTGTATGCGGACAGAGAGAACTGTCCCGCTACTACAGAGCGTGTGCTGAGCTGC	540
QY	688	TGCGACCCGGAAGCATCTTGCGCGCTCTCAGAGTCTGTGGCGCGGAGGTGCTGCAAC	747
Db	541	TGCGACCCGGAAGCATCTTGCGCGCTCTCAGAGTCTGTGGCGCGGAGGTGCTGCAAC	600
QY	748	CTCCGAAAAGGAGCGGTGGCGCGCGAGTGTGCGAAAACCTAAACGAACGATCCGGCGGG	807
Db	601	CTCCGAAAAGGAGCGGTGGCGCGCGAGTGTGCGAAAACCTAAACGAACGATCCGGCGGG	660
QY	808	ACGTACAGGTCTACATCAGCCCTTGCGCCCTGGCGAGTGAAGTCACTTTGGCTTCAAGA	867
Db	661	ACGTACAGGTCTACATCAGCCCTTGCGCCCTGGCGAGTGAAGTCACTTTGGCTTCAAGA	719
QY	868	TCTAAGGCTGGCGCCCTTAGTGAAGTGGAGCTGAG	899
Db	720	TCTAAGGCTGGCGCCCTTAGTGAAGTGGAGCTGAG	750

RESULT 9					
LOCUS	BE796570				
DEFINITION	BE796570	772 bp	mRNA	linear	EST 20-SEP-2000
ACCSSION	601592247	1	NIH_MGC_7	Homo sapiens cDNA clone IMAGE:3946546	5',
VERSION	BE796570				
KEYWORDS	BE796570.1	GI:10217768			
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
REFERENCE	1 (bases 1 to 772)				
AUTHORS	NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .				







Db 62 CGCAGCTGGGCGCGCCTTCCGCACTGGCCTTCTTCTGGGAGGCGGTGCCCATGAC 121  
 Qy 199 GAGGCGCGGAGAGAGAGTCCCTGCTTCCCTCCGAGAGACCGCCTGTTGGAGTATCTTC 258  
 Db 122 GAGGCGCGGAGAGAGAGTCCCTGCTTCCCTCCGAGAGACCGCCTGTTGGAGTATCTTC 181  
 Qy 259 TGAGCGCGTCCATGGGAGAGACCGCGGCTGCGAAGCTTGAAGCTTGACCTTGAGC 318  
 Db 182 TGAGCGCGTCCATGGGAGAGACCGCGGCTGCGAAGCTTGAAGCTTGACCTTGAGC 241  
 Qy 319 AGCGCAGAGGAGGATTTCTATGATGACCTTGAGAGCA-GGCCCACTTGGCCCACTGAC- 376  
 Db 242 AGCGCAGAGGAGGATTTCTATGATGACCTTGAGAGCACTGACCTTGGCCCACTGAC- 301  
 Qy 377 GCGGCTATCCAGGCGCAAGAGGCGCTGAGACCTTGAGCACTTTCAC-GGGCTACTCCGCGC 435  
 Db 302 GCGGCTATCCAGGCGCAAGAGGCGCTGAGACCTTGAGCACTTTCACCTGAGCTACTCCGCGC 361  
 Qy 436 TGGGCTTGGCGCTGGC-GCTGCGCGCGAGACGGGCGCGTGTGACCTTGAGAGTGAACGCG 494  
 Db 362 TGGGCTTGGCGCTGGCCTGCTGCTGCGCGAGACGGGCGCGTGTGACCTTGAGAGTGAACGCG 421  
 Qy 495 CAGCGCGCGAGCTGGGACGGCCTGTTGAGAGGAGGCGGAGGCGGAGCAAGATGAC 554  
 Db 422 CAGCGCGCGAGCTGGGACGGCCTGTTGAGAGGAGGCGGAGGCGGAGCAAGATGAC 481  
 Qy 555 CTCCGCTGAAAGCGCGCTTGGAGACCTTGAGACGAGCTGCGGCGGCGAGGCGGCG 614  
 Db 482 CTCCGCTGAAAGCGCGCTTGGAGACCTTGAGACGAGCTGCGGCGGCGAGGCGGCG 541  
 Qy 615 ACCCTGACGTCGCGCGGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 673  
 Db 542 ACCCTGACGTCGCGCGGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601  
 Qy 674 CTGCTGACGTCGTCGAGACCGCGAGGCA-TCCCTGCGCGTCTGAGAGTCTGTTGGGCGG 732  
 Db 602 CTGCTGACGTCGTCGAGACCGCGAGGCACTTCTGCGCGTCTGAGAGTCTGTTGGGCGG 661  
 Qy 733 GGAAGGTGTCGCACTCCGAAAGGGGAC-GTGGCGCGCAGAGTGTGTGCGAAACTTAAC 791  
 Db 662 GGAAGGTGTCGCACTCCGAAAGGGGACGAGTGTGTGTGCGAAACTTAAC 721  
 Qy 792 GAAAGCATCCGCGGAGGAG-CGTCAAGGTCTTCAATGAGCTTCTGCGCGTGGGCGAGTGG 847  
 Db 722 GAAAGCATCCGCGGAGGAGCAGTCAAGGTCTTCAATGAGCTTCTGCGCGTGGGCGAGTGG 778  
 RESULT 11  
 B0931589 934 bp mRNA linear EST 21-AUG-2002  
 LOCUS AGENCOURT\_8795361 NIH\_MGC\_101 Homo sapiens cDNA clone IMAGE:6427775  
 DEFINITION 5', mRNA sequence.  
 ACCESSION B0931589  
 VERSION B0931589.1 GI:22346620  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
 AUTHORS 1 (bases 1 to 934)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph. D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM2613 row: e column: 24

High quality sequence stop: 532.  
 FEATURES  
 Location/Qualifiers  
 1. 934  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6427775"  
 /issue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_101"  
 /note="Organ: Lung; Vector: pOTB7, site 1: EcoRI; site 2:  
 XhoI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAAGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

Query Match 67.4%; Score 666.4; DB 13; Length 934;  
 Best Local Similarity 95.8%; Pred. No. 7.3e-73;  
 Matches 751; Conservative 0; Mismatches 21; Indels 12; Gaps 6;

Qy 87 ACCAGAGCGGTGCGCGCTCTCCGTCGCGCGCGCGCTGCGCGCGCGCACTG 146  
 Db 1 ACCAGAGCGGTGCGCGCTCTCCGTCGCGCGCGCGCTGCGCGCGCGCACTG 60  
 Qy 147 GGGCGCGCTTCCGCACTGGCTCTTCTGAGAGCGGAGTCCCGCCCATGAGCGCGG 206  
 Db 61 GGGCGCGCTTCCGCACTGGCTCTTCTGAGAGCGGAGTCCCGCCCATGAGCGCGG 120  
 Qy 207 CGAGAGAGTGGCTGCTTCCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 266  
 Db 121 CGAGAGAGTGGCTGCTTCCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 Qy 267 TCCATGCGGAGAGACCGCGCGCTGCGAAGGCTGAGAGTGTGACCCCGAGAGCGGAG 326  
 Db 181 TCCATGCGGAGAGACCGCGCGCTGCGAAGGCTGAGAGTGTGACCCCGAGAGCGGAG 236  
 Qy 327 GGGGATTTCTATGATGACCTTCCGAGAGAGCGCCAGCTTGGCCAACTGGCGGCTCATC 386  
 Db 237 GGGGATTTCTATGATGACCTTCCGAGAGAGCGCCAGCTTGGCCAACTGGCGGCTCATC 296  
 Qy 387 CAGGCGCAAGAGCGGCTGAGACCTTGGGCACTTACCGGCTACTCCGCTTGGCC 446  
 Db 297 CAGGCGCAAGAGCGGCTGAGACCTTGGGCACTTACCGGCTACTCCGCTTGGCC 356  
 Qy 447 CTGGCGCTGCGCGCGAGCGCGCTGTGACCTTGGAGAGTGTGAGCGAGCGCGCGAG 506  
 Db 357 CTGGCGCTGCGCGCGAGCGCGCTGTGACCTTGGAGAGTGTGAGCGAGCGCGCGAG 416  
 Qy 507 CTGGGAGCGGCGCTTGTGAGAGCGCGAGCGGAGCGAGCAAGATGACCTCCGGCTGAAG 566  
 Db 417 CTGGGAGCGGCGCTTGTGAGAGCGCGAGCGGAGCGAGCAAGATGACCTCCGGCTGAAG 476  
 Qy 567 CCGGCTTGGAGAGCCCTGAGAGAGCTGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 626  
 Db 477 CCGGCTTGGAGAGCCCTGAGAGAGCTGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 536  
 Qy 627 GCGGTGTGTGATGCGGAGCAAGAGAGACTGCTCCGCTACTAGAGCGCTGCTGAGCTG 686  
 Db 537 GCGGTGTGTGATGCGGAGCAAGAGAGACTGCTCCGCTACTAGAGCGCTGCTGAGCTG 596  
 Qy 687 CTGGCAACCGGAGGAGATCTCCGCTCTCAAGTCTGTGTGCGGAGGAGTGTGCA 746  
 Db 597 CTGGCAACCGGAGGAGATCTCCGCTCTCAAGTCTGTGTGCGGAGGAGTGTGCA 656  
 Qy 747 CTTCCGAAAGGAGAGT--GGCGCGAGGTGTGT--CGAAGCTTAAAGGAGGCG--ATCC 801  
 Db 657 CTTCCGAAAGGAGAGT 716  
 Qy 802 GGGGAGAGTCA-GGGTCTACATGAGCTCTG--CCCTGGGCGAGTGAAGTCACTCACTTGG 858

Db 717 GGGGGAGCTCAAGGCTTACATCAGCTCCGCGCCCTGCGGAGGAGACTACCTTGG 776  
 QY 859 CCTT 862  
 Db 777 CCTT 780

## RESULT 12

LOCUS B0072503 1065 bp mRNA linear EST 02-APR-2002  
 DEFINITION AGENCOURT 6838941 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5761696  
 5' mRNA sequence.

ACCESSION B0072503  
 VERSION B0072503.1 GI:19901549  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapdb-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: LLM12810 row: 1 column: 17  
 High quality sequence stop: 489.

## FEATURES

source location/Qualifiers  
 1..1065

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5761696"  
 /lab\_host="MDH10B"  
 /clone\_lib="NIH\_MGC\_122"  
 /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;  
 Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source  
 anonymous pool of 24 week female lung, 16 week female  
 spleen, and 20-22 week male spleens. Library is oligo-dT  
 primed and directionally cloned (EcoRV site is destroyed  
 upon cloning). Average insert size 1.4 kb, insert size  
 range 1-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 026. Note:  
 this is a NIH\_MGC Library."

## ORIGIN

Query Match 63.7% Score 630; DB 13; Length 1065;  
 Best Local Similarity 96.7% Pred. No. 2,1e-68;  
 Matches 675; Conservative 0; Mismatches 20; Indels 3; Gaps 3;

QY 1 GCGGGCCCGAGTCCAGAGCTTCCAGAGCTCCAGCTCAGGTAGCTGACAGCC 60  
 Db 1 GCGGGCCCGAGTCCAGAGCTTCCAGAGCTCCAGCTCAGGTAGCTGACAGCC 60  
 QY 61 TCCGCGCGCTCTGCGCGCGCATGACCCAGCGGTGCGCGCTCTCGTGCCCGCG 120  
 Db 61 TCCGCGCGCTCTGCGCGCGCATGACCCAGCGGTGCGCGCTCTCGTGCCCGCG 120  
 QY 121 CGCTGCGCTGAGGCTCAGCGCACTGAGCGCGCGCTTCCGCACTGCGCTTCTTCTGGGA 180  
 Db 121 CGCTGCGCTGAGGCTCAGCGCACTGAGCGCGCGCTTCCGCACTGCGCTTCTTCTGGGA 180  
 QY 181 GCGCGTCCCGCCCTGCGCGCGCATGAGCTGCTGCTTCCCGGAGGAGCAGCC 240

Db 181 GCGCGTCCCGCCCTGCGCGCGCATGAGCTGCTGCTTCCCGGAGGAGCAGCC 240  
 QY 241 GCGTGGCAGTACTTCTGAGCCGCTTCATGCGGAGACCCCGCGCTGCGAACTTGA 300  
 Db 241 GCGTGGCAGTACTTCTGAGCCGCTTCATGCGGAGACCCCGCGCTGCGAACTTGA 300  
 QY 301 GAGTGTGACCTCTGAGAGCGCGAGGAGGATTCATGATGACCTGCGAGAGCCAGC 360  
 Db 301 GAGTGTGACCTCTGAGAGCGCGAGGAGGATTCATGATGACCTGCGAGAGCCAGC 360  
 QY 361 TCTTGGCCAACTGAGCGCGCTCTATCCAGGCGCAAGAGCGCTGAGCTGAGCACTTCA 420  
 Db 361 TCTTGGCCAACTGAGCGCGCTCTATCCAGGCGCAAGAGCGCTGAGCTGAGCACTTCA 420  
 QY 421 CGGCTACTCGCGCCCTGAGCGCGCTGAGCGCGCTGAGCGCGCGCTGAGCGCGCT 480  
 Db 421 CGGCTACTCGCGCCCTGAGCGCGCTGAGCGCGCTGAGCGCGCGCTGAGCGCGCT 480  
 QY 481 GCGAGGTGAGCGCGCAGCGCGCGCGCTGAGCGCGCGCTGAGCGCGCGCGCG 540  
 Db 481 GCGAGGTGAGCGCGCAGCGCGCGCGCTGAGCGCGCGCTGAGCGCGCGCGCG 540  
 QY 541 AGCAAGATGACCTTCCGCTGAGCGCGCGCTTGGAGA CCTGAGCAGCTGCGCG 599  
 Db 541 AGCAAGATGACCTTCCGCTGAGCGCGCGCTTGGAGA CCTGAGCAGCTGCGCG 600  
 QY 600 GCGGGGAGCGCGCGCGCGCTGAGCGCGCGCTGAGCGCGCGCGCGCGCGCT 657  
 Db 601 GAGCGGAGCGCGCGCGCGCGCGCTTGAAGTGCAGCGGAGGAGCGAGCAAGAAAATGCT 660  
 QY 658 CCGCTTACTGAGAGCGCTGCGCTGAGCGCGCTGAGCGCGCGCGCGCGCGCT 695  
 Db 661 CCGCTTACTGAGAGCGCGCTGAGCGCGCGCTGAGCGCGCGCGCGCGCGCT 698

## RESULT 13

LOCUS BM768764 625 bp mRNA linear EST 04-MAR-2002  
 DEFINITION K-EST0051789 S14K402 Homo sapiens cDNA clone S14K402-11-A01 5',  
 mRNA sequence.

ACCESSION BM768764  
 VERSION BM768764.1 GI:19098379  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
 Kim, Y.S.

TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Deon-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 11 row: A column: 01  
 High quality sequence stop: 625.

## FEATURES

source location/Qualifiers  
 1..625

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S14K402-11-A01"  
 /cell\_line="K402"  
 /lab\_host="Top10F"  
 /clone\_lib="S14K402"  
 /note="Organ: Stomach; Vector: pTZ198Rpl; Site\_1: EcoRI;  
 Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with



Db 421 CGAGCCCCCGAGCTGGAGCGG-CCTGTGAGGAGCGAGCGGAGCAGACAGATCG 479

QY 553 ACCTCCGGCTGAGACCCGCGCTTGGAGACCTGGAGAGCTGCTGGGGGGGAGAGCCG 612

Db 480 ACCTCCGGCTGAGACCCGCGCTTGGAGACCTGGAGAGCTGCTGGGGGGGAGAGCCG 539

QY 613 GCACCTTCGACGCTGGCGCTGTGTGATGCGGACAGAGAGAACTGCTCCGCTACTACAGC 672

Db 540 GCACCTTCGACGCTGGCGCTGTGTGATGCGGACAGAGAGAACTGCTCCGCTACTACAGC 596

QY 673 GCTGCTGAGAGCTGCTGCGGACCCGAGAGCACTCTGCGCCCTCTAGAGTCTGTGGCGG 732

Db 597 GCTGCTGAGAGCTGCTGCGGACCCGAGAGCACTCTGCGCCCTCTAGAGTCTGTGGCGG 654

QY 733 GGAAGGTGCTGCA-CTCCGAAAGGGAGAGCTGGCGCCGAGTGTGCGCAAACTTAAAG 792

Db 655 GGAAGGTGCTGCA-CTCCGAAAGGGAGAGCTGGCGCCGAGTGTGCGCAAACTTAAAG 710

QY 793 AAGCATCCGCGGAGCTGAGAGTCTATCATCAGCTCTGCGCCCTGAG 841

Db 711 AAGCATCCGCGGAGCTGAGAGTCTATCATCAGCTCTGCGCCCTGAG 758

RESULT 15

AK054334

LOCUS 2243 bp mRNA linear HTC 20-SEP-2003

DEFINITION Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330016L21 product:weakly similar to PUTATIVE O-METHYLTRANSFERASE (FRAGMENT) [Rhodothermus marinus], full insert sequence.

ACCESSION AK054334

VERSION AK054334.1 GI:26344156

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Carninci, P. and Hayashizaki, Y.

JOURNAL High-efficiency full-length cDNA cloning

MEDLINE 99279253

PMID 10349636

REFERENCE Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL 2049374

MEDLINE 11042159

PMID 11042159

REFERENCE Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, Y., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M., Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunaga, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubara, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL 20530913

MEDLINE 11076861

PMID 11076861

REFERENCE The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

JOURNAL 5

MEDLINE

PMID

REFERENCE The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation

JOURNAL of 60,770 full-length cDNAs

REFERENCE Nature 420, 563-573 (2002)

AUTHORS Adachi, J., Aizawa, K., Akiyama, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imomani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanakawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohse, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, K., Takahashi, F., Takaku, Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUN-2001) Yoshitake Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://phantom.gsc.riken.go.jp/

FEATURES

source

1..2243

Location/Qualifiers

1..2243

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="PANTOM:DB:E330016L21"

/db\_xref="MG1:2426440"

/db\_xref="taxon:10090"

/clone="E330016L21"

/sex="female"

/tissue type="ovary"

/clone\_id="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="2 days pregnant adult"

40..828

/note="unnamed protein product; putative weakly similar to PUTATIVE O-METHYLTRANSFERASE (FRAGMENT) [Rhodothermus marinus] (SPR109XCB3, evidence: PASTY, 55.5%ID, 97.3%length, match=331)"

/codon\_start=1

/protein\_id="BAC35735.1"

/db\_xref="GI:26344157"

/translation="MAQPVPRISIPALALGSAALGAFAATGLIGKRWPPGSRROE RLIPEDNPIMQYILSRSHREHLSIRLTLEOPGDSMTCEQALANTLRLIK AKRALDGTFTGSAALALALBEAGVVCVDAEPKGRPMKQAEVOKDRLR OPLQTDLELALGEAGETPDIAVDADKENCYVERCCLIRPGGLAVLRVLRRE VLDQRPQNKVEECVRNINERILADARYISILPLDDSLAFKI"

2223..2228

/note="putative"

2243

/note="putative"

ORIGIN

polya\_signal

polya\_site

Query Match 61.6%; Score 609.2; DB 11; Length 2243;

Best Local Similarity 80.9%; Pred. No. 6.2e-66;

Matches 723; Conservative 0; Mismatches 168; Indels 3; Gaps 1;

QY 63 CCGCGGCTCTGCGCGGCGGATGACCCAGCGGCTGCGGCTCTGCGCGGCGGCG 122

Db 19 CCGCGGCTCTGCGCGGCGGATGACCCAGCGGCTGCGGCTCTGCGCGGCGGCG 78

QY 123 CTGCGCTGGGCTCAAGCGGCACTGGGCGCGGCTTGGCCACTGCGCTTCTGCGGAGG 182





OY	408	CTGGGACACTTCAACGGGCTACTCCCGCCTTAGCCTTGACCCTTGAGCGCTGCCGCGGAACGGG	467
Dp	546	CTGGGACCTTCACAGGCTACTCCGCCTTAGCCTTGACCCTTGAGCGCTGCCGCGGAACGGG	487
OY	468	CAGGTGTATGACTGCGCATGTGACACGCGCACGCCCGGAGATGTGGACCGCCCCCTGTGAGG	527
Dp	486	CGGTGTGTACTCTGCGAGGTGACACGCGCACGCCCGGAGNTGTGACCGCCCCCTGTGAGG	427
OY	528	CAGGCGGAGCGGAGACAAGAATCGACTTCGCGCTGAAGCCCGCCTTTGAGAACCTTGAGAC	587
Dp	426	CAGGCGGAGCGGAGACAAGAATCGACTTCGCGCTGAAGCCCGCCTTTGAGAACCTTGAGAC	367
OY	588	GACCTGTGGCGGGCGGCGGAGGCGGACCTTTGACGTGGCGGTGTGATGCGACAAG	647
Dp	366	GACCTGTGGCGGGCGGCGGAGGCGGACCTTTGACGTGGCGGTGTGATGCGACAAG	307
OY	648	GAGAACTGCTCCGCTACTACTACGAGCGCTGCTGACACTGCTGTGGACCCGAGCACTTCTC	707
Dp	306	GAGAACTGCTCCGCTACTACTACGAGCGCTGCTGACACTGCTGTGGACCCGAGCACTTCTC	247
OY	708	GCCGTCTTACAGATCCTGTGGCGGGGAAAGTGTCTCAACTCCGAAAAGGAGACGTGGCG	767
Dp	246	GCCGTCTTACAGATCCTGTGGCGGGGAAAGTGTCTCAACTCCGAAAAGGAGACGTGGCG	187
OY	768	GCGAGTGTGTCCGAAACCTTAACCAAACCAACCAATCCGCGGGAAGTCAAGGTCTTACATCAGC	827
Dp	186	GCGAGTGTGTGTGAGAAACCTTAACCAAACCAACCAATCCGCGGGAAGTCAAGGTCTTACATCAGC	127
OY	828	CTCTGACCCCTGGGCGAATGGAATCACTACCTTGAGCCTTCAAGATTTAGAGGCTGAGCCCTAGTG	887
Dp	126	CTCTGACCCCTGGGCGAATGGAATCACTACCTTGAGCCTTCAAGATTTAGAGGCTGAGCCCTAGTG	67
OY	888	AGTGGGCTTCGAGGAGAGGCTTGCCTTGGAAACCCACAGAAATTGACCTTGAGTTTTAAATTCG	947
Dp	66	AGTGGGCTTCGAGGAGAGGCTTGCCTTGGAAACCCACAGAAATTGACCTTGAGTTTTAAATTCG	7
OY	948	AAAAATA 953	
Dp	6	AAAAATA 1	
RESULT 18			
LOCUS	BY707694		
DEFINITION	musculus full-length enriched, 10 day old male pancreas Mus	mRNA linear EST 16-DEC-2002	
VERSION	BY707694.1	GI:27118872	
KEYWORDS	Mus musculus (house mouse)		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
AUTHORS	Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,		
	Kikaido,I., Otsu,N., Saito,R., Suzuki,H., Yamana,K.I.,		
	Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,		
	Schomach,C., Gotohori,T., Baldarelli,R., Hill,D.P., Bull,C.,		
	Hume,D.A., Quackembush,J., Schirral,L.M., Kanagin,A., Matsuda,H.,		
	Batalov,S., Beisel,K.W., Blake,U.A., Bradt,D., Busnic,V.		
	Chochia,C., Corpani,L.E., Cousins,S., Della,E., Dragani,T.A.,		
	Pletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,		
	Gariboldi,M., Glisi,C., Godzik,A., Gough,U., Grimmond,S.,		
	Gustinghi,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,		
	Kawajiri,H., Kawasawa,Y., Kedzierzki,R.M., King,B.L., Konagaya,A.,		
	Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,		
	Maltsis,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,		
	Nunata,K., Okido,T., Pavan,W.U., Pertea,G., Pesole,G.,		
	Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,		
	Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Rungwald,M.,		
	Sandelin,A., Schneider,C., Sample,C.A., Setou,M., Shimada,K.,		
	Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,		
	Verrardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,		

	Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
	Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
	Hayatsu, N., Hirozane-Kishikawa, T., Komno, H., Nakamura, M.,
	Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
	Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.
	Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Saeki, D., Shibata, K.,
	Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
	Rogers, J., Birney, E., and Hayashizaki, Y.
TITLE	Analysis of the mouse transcriptome based on functional annotation
JOURNAL	of 60,770 full-length cDNAs
MEDLINE	Nature 420, 563-573 (2002)
PUBMED	22354683
COMMENT	12466851
	Contact: Yoshihide Hayashizaki
	Laboratory for Genome Exploration Research Group, RIKEN Genomic
	Sciences Center(GSC), Yokohama Institute
	The Institute of Physical and Chemical Research (RIKEN)
	1-7-22 Shihori-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
	Tel.: 81-45-503-9222
	Fax: 81-45-503-9216
	Email: genome-res@gsc.riken.go.jp,
	url:http://genome.gsc.riken.go.jp/
	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
	Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Horii, F.,
	Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,
	Komoda, S., Komno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, N.,
	Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
	Sano, H., Saeki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
	Takeda, Y., Waki, K., Watabiki, A., Muramatsu, M. and Hayashizaki, Y.
	Direct Submission
	Computational Analysis of Full-Length Mouse cDNAs Compared with
	Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
	Normalization and subtraction of cap-trapper-selected cDNAs to
	prepare full-length cDNA libraries for rapid discovery of new
	genes. Genome Res. 10 (10), 1617-1630 (2000)
	RIKEN integrated sequence analysis (RISA) system-384-format
	sequencing pipeline with 384 multicapillary sequencer. Genome Res
	10 (11), 1757-1771 (2000)
	Computer-based methods for the mouse full-length cDNA
	nonredundant cDNA library. Genome Res. 11 (2), 261-269 (2001)
	cDNA library was prepared and sequenced in Mouse Genome
	Encyclopedia Project of Genome Exploration Research Group in RIKEN
	Genomic Sciences Center and Genome Science Laboratory in RIKEN.
	Division of Experimental Animal Research in Riken contributed to
	prepare mouse tissues.
	Please visit our web site ( <a href="http://genome.gsc.riken.go.jp">http://genome.gsc.riken.go.jp</a> ) for
	further details.
FEATURES	Location/Qualifiers
source	1..910
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"
	/clone="181003OM08"
	/sex="male"
	/tissue_type="pancreas"
	/dev_stage="10 day old"
	/clone_lib="RIKEN full-length enriched, 10 day old male
	pancreas"
ORIGIN	
Query Match	59.4%; Score 587.8; DB 13; Length 910;
Best Local Similarity	82.3%; Pred. No. 3.4e-63;
Matches 687; Conservative 0; Mismatches 147; Indels 1; Gaps 1	
D6	63 CGCGGCTCTGTCGCCGGCCAGACCAAGCCGATGCCCGGCTCTCGTGCAGCGCG 122
OY	13 CGCGCGCTGTAAACTGACCATGCTCAGCGCCGTCCTCGCTATTCTATCCAGCGGA 72
OY	123 CTGGGCTCTGGGCTCAGCGCACATGGGGCGCGCTTGCGCATGGCGCTTTCTTGGGAGG 182
D6	73 CTGGCGCTGGGCTCGCGCGCTGGGCG-CGCTTGCGTACTAGTGTCTTGGCTGGGAAA 131





RNA sequence.  
 B339315  
 VERSION B339315.1 GI:13145753  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1122)  
 NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgsbds-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 Cloning Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: LNCMI256 row: n column: 02  
 High quality sequence stop: 730.

FEATURES  
 source  
 Location/Qualifiers  
 1..1122  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:455585"  
 /tissue\_type="leiomyosarcoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH MGC 46"  
 /note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

## ORIGIN

Query Match 59.2%; Score 585.6; DB 12; Length 1122;  
 Best Local Similarity 87.2%; Pred. No. 5.9e-63;  
 Matches 806; Conservative 0; Mismatches 94; Indels 24; Gaps 14;

73 CTGCGCGCGGCAAGACCCAGCCGCTGCGGCTCTCCGCGCGCGCTGCGCGCTGCG 132  
 2 CTGCGCGCGGCAAGACCCAGCCGCTGCGGCTCTCCGCGCGCGCTGCGCGCTGCG 61  
 133 GCTCAGCGCGC-ACTGAGGCGCGCGCTTCCGCACTGGGCTCTCTGCGGAGGCGGTCGCC 191  
 62 GCTCAGCGCGCTACTGCGCGCGCGCTTCCGCACTGGGCTCTCTGCGGAGGCGGTCGCC 121  
 192 CCATGCGAGGCGCGC-GCGAGAGAGTGCCTGCTTCCCGCGAGAGAGCGGCTGAGGA 250  
 122 CCATGCGAGGCGCGTCCGAGAGAGAGTGCCTGCTTCCCGCGAGAGAGCGGCTGAGGA 181  
 251 GATATCTTGAAGCGCGCTCCATGCGGAGAGACCGCGC-GCTGCGAAGCTTGAAGGCTGCTGA 309  
 182 GATATCTTGAAGCGCGCTCCATGCGGAGAGACCGCGAGCGCTGCGAAGCTTGAAGGCTGCTGA 241  
 310 CCTGAGAGAGCGCGAGGAGGATTTCTA-TGATGACTTGGAGAGAGCGCGAGCTTGGGCG 368  
 242 CCTGAGAGAGCGCGAGGAGGATTTCTAAGTATGATGATCTTGGAGAGAGCGCGAGCTTGGGCG 301  
 369 AACCTG---GCGGCGCTCATTCAGAGCGCAAGAGCGCTGAGCTGAGGCACTTTCAGCGGCG 425  
 302 AACCTGCGCGAGCTCATTCAGAGCGCAAGAGCGCTGAGCTGAGGCACTTTCAGCGGCG 361  
 426 TACTCGCGCTGCG-CCCTGCGCGCTGCGCTGCGCGCGAGCGGCGGCTGATGACTGCGCA 484

Db 362 TACTCGCGCTGCGCTGCGCGCTGCGCTGCGCGAGCGGCGGCTGATGACTGCGCA 421  
 QY 485 GGTGAGCGCGAGCGCGCGCGCGAGCTGAGAGCGCGCGCTGAGAGAGCGCGCA-GGCGAGCG 543  
 Db 422 GGTGAGCGCGAGCGCGCGCGCGAGCTGAGAGCGCGCGCTGAGAGAGCGCGCAAGCGAGCG 481  
 QY 544 ACAAGATGACCTTCCGCTGAAGCGCGCTTGGAGACCTTGGAGAGCGAGCTGCGCGCGG 603  
 Db 482 ACAAGATGACCTTCCGCTGAAGCGCGAGCTTGGAGAGCGAGCTGCGCGCGG 541  
 QY 604 GCGAGCGCGCGCACTTGGAGAGCGTGGCGGATGAGTGGAGCAAGAGAGAGCTGCGCGCT 663  
 Db 542 GCGAGCGCGCGCACTTGGAGAGCGTGGCGGATGAGTGGAGCAAGAGAGAGCTGCGCGCT 601  
 QY 664 ACTAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723  
 Db 602 ACTAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658  
 QY 724 TGTGCGCGCGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783  
 Db 659 TGTGCGCGCGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 718  
 QY 784 ACCTAAAGCAAC-GCATCCGCGCGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 841  
 Db 719 ACCTAAAGCAACGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 778  
 QY 842 CGATGAGCTACCTTGGCGCTTCAAGATCTAGAGCTGCGCGCTGAGTGTGCTGAGAGG 901  
 Db 779 CGAAGAGCTCA-CTTGGCGCTTCAAGATCTAGAGCTGCGCGCTGAGTGTGCTGAGAG 832  
 QY 902 AGGATGCGCGGAGAGCGCGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961  
 Db 833 AGGATGCGCGGAGAGCGCGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 888  
 QY 962 CTGCGGACACAAAAA 985  
 Db 889 GGGCGCGGAGAAAAA 912

## RESULT 21

BM769685 580 bp mRNA linear EST 04-MAR-2002  
 LOCUS K-EST0052976 S14K402 Homo sapiens cDNA clone S14K402-22-C12 5',  
 DEFINITION mRNA sequence.

ACCESSION BM769685  
 VERSION BM769685.1 GI:19099300

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Kim,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Oh,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS

Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.krrib.re.kr  
 Plate: 22 row: C column: 12  
 High quality sequence stop: 580.  
 Location/Qualifiers  
 1..580  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S14K402-22-C12"  
 /cell\_line="K402"

## FEATURES

source



enrichment and has not undergone amplification. Library was constructed by Dr. Claudio Schneider (INCI-B-Area Science Park, Trieste, Italy). "

Query Match	58.5%;	Score 578.4;	DB 9;	Length 581;
Best Local Similarity	99.7%;	Pred. No. 5.4e-62;		
Matches 579; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

QY	43	ACGTACCTGTACATCGCTCCCGCGGCTCCGACCGGCATATACCAAGCGGTGGCC	102
Db	1	ACGTACCTGTACATCGCTCCCGCGGCTCCGACCGGCATATACCAAGCGGTGGCC	60
QY	103	GAGCTCTCGGTGCGCGCCGCGTGGCGCTCAGGCGCATGAGCGCCGCTTGCCA	162
Db	61	GAGCTCTCGGTGCGCGCGCTGGCGCTGAGCGGCATGAGCGCCGCTTGCCA	120
QY	163	CTGGGCTCTTCTTGGGGAGGGGGTGCCTCCCATGCGAGGCGCGAGAGCATGTGCTGC	222
Db	121	CTGGGCTCTTCTTGGGGAGGGGGTGCCTCCCATGCGAGGCGCGAGCATGTGCTGC	180
QY	223	TTCCCCCGAGAGACAGCGCCGTGGGAGATATCTTGGAGCGCGCTCCATGCGGAGACAC	282
Db	181	TTCCCCCGAGAGACAGCGCCGTGGGAGATATCTTGGAGCGCGCTCCATGCGGAGACAC	240
QY	283	CGGCGCTGCGAAGCTTGAAGCTGTGTGAACCTGAGACAGCGCAGGGGATTTCTATGATGA	342
Db	241	CGGCGCTGCGAAGCTTGAAGCTGTGTGAACCTGAGACAGCGCAGGGGATTTCTATGATGA	300
QY	343	CCTGCGAGCAGGCCCATAGCTTTGGCCAACTGGCGGCTCATCCAGGCCAAGAAAGCGGC	402
Db	301	CCTGCGAGCAGGCCCATAGCTTTGGCCAACTGGCGGCTCATCCAGGCCAAGAAAGCGGC	360
QY	403	TGGAACCTGGGCACTTACCGGGACTACCTCCGCGCTGGGCGCTGGCGCGTGGCCGCGG	462
Db	361	TGGAACCTGGGCACTTACCGGGACTACCTCCGCGCTGGGCGCTGGCGCGTGGCCGCGG	420
QY	463	ACGGGCGGTGTGTGACCTGCGAGGTGAGACGCGACGCCCGGAGCTGGGAACGCGCCCTGT	522
Db	421	ACGGGCGGTGTGTGACCTGCGAGGTGAGACGCGACGCCCGGAGCTGGGAACGCGCCCTGT	480
QY	523	GGAGGCGAGCCGAGGCGGAGACAAAGATCGACTTCGAGCTTGAAGCCCGCCTTGGAGACCC	582
Db	481	GGAGGCGAGCCGAGGCGGAGACAAAGATCGACTTCGAGCTTGAAGCCCGCCTTGGAGACCC	540
QY	583	TGAGCGAGCTGTGGCGGCGGCGGAGAGCGCGACCTTTGAC	623
Db	541	TGAGCGAGCTGTGGCGGCGGCGGAGAGCGCGACCTTTGAC	581

**Estadística**

**source**

WashU-NCI Human EST Project  
Unpublished (1997)  
Other ESTs: au59603.x1  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
This clone is available royalty-free through LINT; contact the  
IMaGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Possible reversed clone; similarity on wrong strand  
Seq primer: -40RP from Gibco  
High quality sequence stop: 368

TITLE	Wasnu-NCI human EST Project
JOURNAL	Unpublished (1997)
COMMENT	Other_ESTS: au59e03.x1

Db	573	AAGTCTGTGGCGCGGGAAAGTGCTGCAACCTCCGAAAGGGACGTGGCGGGCCGAGTGT	632
Qy	777	GTCCGAACCTTAAACGAACGACATTCGGCGGGACGTCAAGGATCTTACATCAGCTCTGTGCC	836
Db	633	GTCCGAACCTTAAACGAACGACATTCGGCGGGACGTCAAGGATCTTAAATCAGCTCTGTGCC	652
Qy	837	CTGGGCGATGGAGTCAAC--TTGGGCTTCAAGATCTTGGGGTGGCCCCCTAGTGAATGGG--C	894
Db	693	CTGGGCGATGGAGTCTCCCTTTGGCTTTTCAAGATCTTGGGGCTGGCCCCCTAGTGAATGGGCT	752
Qy	895	TCGAGGAGAGGTTGCTCTGGGAA--CCCCAGGATTTGACCTGATTTTAAATTCGAAAT	952
Db	753	TCAGGGAGGTTGCCCCGGGAAACCCACAGAAATGACCCGGGAGTTTAAATTCGAAAT	812
Qy	953	AAAGTGGGG	961
Db	813	CAAGGGGGG	821

RESULT 23	AI929359	AI929359	581 bp	mRNA	linear	EST 23-AUG-1999
LOCUS	AI929359					
DEFINITION	IMAGE03.Y1 Schneider fetal brain 00004 Homo sapiens CDNA clone IMAGE:511068.5', similar to FR:085769 085769 HYPOTHETICAL 24.8 KD PROTEIN. ; mRNA sequence.					
ACCESSION	AI929359					

```

FEATURES
SOURCE
Location/Qualifiers
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2519068"
/sex="male"
/tissue_type="frontal lobe"
/dev_stage="5 months post-conception"
/lab_host="DH10B"
/clone_lib="Schneider fetal brain 00004"
/note="Organ: brain; Vector: pBluescript SK (Stratagene);
Site_1: SstI; Site_2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GAGGAGAGAGAGAGCTCAGAGATCCTTAATTAATTAATCCCCCCCCCCC-3'
and 3' adaptor sequence:
5'-GAGGAGAGAGAGCTCAGATTTTTTTTTTTTT-3'. The library was
size-selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the CAP-trapper method for full-length

```



474 GTGACCTGGAGGTGAGACGGCGACGCCCGGAGCTGGAGCGGCTGTGGAGCAGGCC 533  
Db |TGAAGCTGGAGGTGAGACGGCGACGCCCGGAGCTGGAGCGGCTGTGGAGCAGGCCA 420  
Qy 534 GAGGCGGACCAAGATCGACCTCCCGCTGAGACCCCGCTTGGAGACCTTGGAGCAGCTG 593  
Db 421 GAAGTGGACCAAGATCGACCTCCCGCTGAGACCCCGCTTGGAGACCTTGGAGCAGCTG 480  
Qy 594 CTGGCGGCGGCGGAGCGCGGACCTTGGAGCGGCTGTGGAGCGGCTGTGGAGCGGAC 653  
Db 481 CTAGCGGCGGCGGAGCGCGGACCTTGGAGCGGCTGTGGAGCGGCTGTGGAGCGGAC 540  
Qy 654 TGCTCGGCTTCTAGAGCGGCTGTGGAGCGGCTGTGGAGCGGCTGTGGAGCGGCTGTG 713  
Db 541 TGACCGGCTTCTAGAGCGGCTGTGGAGCGGCTGTGGAGCGGCTGTGGAGCGGCTGTG 600  
Qy 714 CTGAGAGTCTGTGGCGGCGGAGGAGTGTGGAGCGGCTGTGGAGCGGCTGTGGAGCGG 773  
Db 601 CTGAGAGTCTGTGGCGGCGGAGGAGTGTGGAGCGGCTGTGGAGCGGCTGTGGAGCGG 660  
Qy 774 TGTGTGGGAAACCTTAACGAAAGCATCGGCGGAGGAGTGTGGAGCGGCTGTGGAGCGG 833  
Db 661 TGTGTGGGAAACCTTAACGAAAGCATCGGCGGAGGAGTGTGGAGCGGCTGTGGAGCGG 720  
Qy 834 CCCCTGGCGGAGTGG-ACCTCAGCTTGGCTTGAAGTCTGAAGTCTGAAGTCTGAAGTCTG 890  
Db 721 CCCCTGGCGGAGTGG-ACCTCAGCTTGGCTTGAAGTCTGAAGTCTGAAGTCTGAAGTCTG 780  
Qy 891 GAGGCTGAGGAGGAGGAGTGTGGAGCGGCGGAGGAGTGTGGAGCGGCTGTGGAGCGG 947  
Db 781 GAGGCTGAGGAGTGTGGAGCGGCGGAGGAGTGTGGAGCGGCTGTGGAGCGGCTGTGGAG 947  
Qy 948 AAAATTAAGTGGGCTGTGGAGCGGCGGAGGAGTGTGGAGCGGCTGTGGAGCGGCTGTG 840  
Db 841 ACAATTAAGTGTGAGGAGCGGCGGAGGAGTGTGGAGCGGCTGTGGAGCGGCTGTGGAG 882

RESULT 26  
LOCUS AM157252/c 582 bp mRNA linear EST 04-NOV-1999  
DEFINITION au3801.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone  
IMAGE:2783832.3' similar to TR:085769 085769 HYPOHETICAL 24.8 KD  
PROTEIN. ; mRNA sequence.

ACCESSION AM157252  
VERSION AM157252.1 GI:6228653

KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 582)  
AUTHORS Hillier, L., Allen, M., Bowles, J., Dubuque, T., Geisler, G., Jost, S.,  
Kizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,  
Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F.,  
Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.  
WashU-MC human EST Project  
Unpublished (1997)  
CONTACT: Wilson RK

TITLE Washington University School of Medicine  
JOURNAL 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
COMMENT Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
This clone is available royalty-free through INM; contact the  
IMAGE Consortium (info@image.inl.gov) for further information.  
High quality sequence stop: 419.  
Location/Qualifiers  
1.582  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2783832"

/sex="male"  
/tissue\_type="frontal lobe"  
/dev\_stage="5 months post-conception"  
/lab\_host="DH10B"  
/clone\_lib="Schneider fetal brain 00004"  
/note="Organ: brain; Vector: pBluescript SK (Stratagene);  
Site\_1: Sct1; Site\_2: Xho1; Double-stranded cDNA was  
prepared from human fetal brain tissue. 5' and 3'  
adaptors were used in cloning as follows: 5' adaptor  
sequence:  
5'-GAGAGAGAGAGAGCTCAAGATCCTTAATTAATTAATTCCTCCCTCCCTCC-3'  
and 3' adaptor sequence:  
5'-GAGAGAGAGAGCTCAAGATCCTTAATTAATTAATTCCTCCCTCCCTCC-3'. The library was  
size-selected for >0.5 kb inserts and has an average  
insert size estimated at 1.2 kb. This library was  
constructed using the CAP-trapper method for full-length  
enrichment and has not undergone amplification. Library  
was constructed by Dr. Claudio Schneider (LNCIB-Area  
Science Park, Trieste, Italy)."

Query Match 57.1%; Score 564.4; DB 10; Length 582;  
Best Local Similarity 98.1%; Pred. No. 2.9e-60;  
Matches 571; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 372 CTGGCGGCGGCTCATCCAGGCGGAGAGGCGGCTGGAGCTGGAGCGGCTGTGGAGCGG 431  
Db 582 CTGGCGGCGGCTCATCCAGGCGGAGGAGGCGGCTGGAGCTGGAGCGGCTGTGGAGCGG 523  
Qy 432 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 491  
Db 522 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 463  
Qy 492 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 551  
Db 462 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 403  
Qy 552 GACCTCGGCTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 611  
Db 402 GACCTCGGCTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 343  
Qy 612 GGCACCTTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 671  
Db 342 GGCACCTTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 283  
Qy 672 GCGTGGCTTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 731  
Db 282 GCGTGGCTTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 223  
Qy 732 GGGAGAGTCTGCAAGCTCCGAAAGGAGCGTGGCGGCGGCGGCGGCGGCGGCGGCGG 791  
Db 222 GGGAGAGTCTGCAAGCTCCGAAAGGAGCGTGGCGGCGGCGGCGGCGGCGGCGGCGG 163  
Qy 792 GAAAGCATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 851  
Db 162 GAAAGCATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 103  
Qy 852 ACCCTGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 911  
Db 102 ACCCTGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 43  
Qy 912 GGGAGAGTCTGCAAGCTCCGAAAGGAGCGTGGCGGCGGCGGCGGCGGCGGCGGCGG 953  
Db 42 GGGAGAGTCTGCAAGCTCCGAAAGGAGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1

RESULT 27  
LOCUS CB181160 886 bp mRNA linear EST 31-JAN-2003  
DEFINITION AGENCOURT\_11381817 NIH MGC\_164 Mus musculus cDNA clone  
IMAGE:30240746.5', mRNA sequence.  
ACCESSION CB181160  
VERSION CB181160.1 GI:28177885









/clone="IMAGE:30135806"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NH1\_MGC\_135"  
 /note="Vector: PCW5port6.1; Site 1: EcoRV; Site 2: NotI;  
 Normalized full-length enriched library from pooled mouse  
 embryonic limb, maxilla and mandible, day 12.5, 13.5,  
 14.5 and 15.5 (size selected for the 0.5-1 kb fragments)  
 cloned directionally, priming method: Oligo-dT. cDNA  
 enrichment: >1k bp. Average insert size 1.6k bp.  
 Normalization (cot value): 7.5 kb. Priming sequence:  
 5'-GACTAGTTCTAGATCGGAGCGGCCGCC(T)3', tissue contributed  
 by David Rowe. Library constructed by ResGen, Invitrogen  
 Corp."

## ORIGIN

Query Match 54.9%; Score 542.8; DB 14; Length 909;  
 Best Local Similarity 82.6%; Pred. No. 1.1e-57;  
 Matches 634; Conservative 0; Mismatches 132; Indels 2; Gaps 1;

```

63 CCGCGGCTCTCTCCCGGCGCCATGACCCGCGGCTCTCGTGGCCGCGCG
3 CCGCGGCTCTCTCCCGGCGCCATGACCCGCGGCTCTCTATATCCGACCGCA
123 CTGGCCCTGGGCTCAAGCGGCACTGGGCGCGGCTCTCGCCTCTCTCTGGGAGG
63 CTGGCCCTGGGCTCAAGCGGCGGCTGGGCGCGGCTCTCTCTCTGGGAGAA
183 CCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
123 CCGTGGCTCTCATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
243 CTGTGGCGAGTATCTTCTGAGCGGCTCTCATGGGCGGCGGCGGCGGCGG
183 CTGTGGCGAGTATCTTCTGAGCGGCTCTCATGGGCGGCGGCGGCGGCGG
303 CTGTGGCGAGTATCTTCTGAGCGGCTCTCATGGGCGGCGGCGGCGGCGG
243 CTGTGGCGAGTATCTTCTGAGCGGCTCTCATGGGCGGCGGCGGCGGCGG
363 TTGGCGCAACTGGGCGGCGGCTCTCATGGGCGGCGGCGGCGGCGGCGG
303 CTGGCGCAACTGGGCGGCGGCTCTCATGGGCGGCGGCGGCGGCGGCGG
423 GGCTACTCTGGGCGGCGGCTCTCATGGGCGGCGGCGGCGGCGGCGGCGG
363 GGCTACTCTGGGCGGCGGCTCTCATGGGCGGCGGCGGCGGCGGCGGCGG
483 GAGGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
423 GAGGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
543 CACAAGATGACCTTCGGGCTGAAAGCGGCTTGGAGACCTTGGAGACCTT
483 CACAAGATGACCTTCGGGCTGAAAGCGGCTTGGAGACCTTGGAGACCTT
603 GGGCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
543 GGGCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
663 TACTACGAGCGGCTCTCTGAGCGGCTCTGAGACCGGCGGCGGCGGCGG
603 TACTACGAGCGGCTCTCTGAGCGGCTCTGAGACCGGCGGCGGCGGCGG
723 CTGTGGCGGCGGAGAGTGTCTGCAACTTCGAAAGAGGAGCGTGGCGGCG
782 CTGTGGCGGCGGAGAGTGTCTGCAACTTCGAAAGAGGAGCGTGGCGGCG
663 CTGTGGCGGCGGAGAGTGTCTGCAACTTCGAAAGAGGAGCGTGGCGGCG
722 CTGTGGCGGCGGAGAGTGTCTGCAACTTCGAAAGAGGAGCGTGGCGGCG
783 AACTTAAGAGAGCGATTCGGGCGGAGCGTC--AGGCTTATATCAAGCC 828
723 AACTTAAGAGAGCGATTCGGGCGGAGCGTC--AGGCTTATATCAAGCC 770

```

RESULT 32

AI765658/c  
 LOCUS AI765658 572 bp mRNA linear EST 21-BRC-1999  
 DEFINITION w182g12.x1 NCI CGAP Kid12 Homo sapiens cDNA clone IMAGE:2399878 3'  
 similar to TR:085769 085769 HYPOTHETICAL 24.8 KD PROTEIN. ; mRNA  
 sequence.  
 ACCESSION AI765658.1 GI:5232167  
 VERSION AI765658  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 572)  
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 www.bio.lnl.gov/bdnp/image/image.html  
 Insert Length: 974 Std Error: 0.00  
 Seq primer: -400p from Glpco  
 High quality sequence stop: 467.  
 Location/Qualifiers  
 1..572  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2399878"  
 /tissue\_type="2 pooled tumors (clear cell type)"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Kid12"  
 /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with  
 a modified polylinker; Site 1: Not I; Site 2: Eco RI;  
 Plasmid DNA from the normalized library NCI CGAP Kid5 was  
 prepared, and ss circles were made in vitro. Following HAP  
 purification, this DNA was used as tracer in a subtractive  
 hybridization reaction. The driver was PCR-amplified cDNAs  
 from a pool of 5,000 clones made from the same library  
 (clonoids 1323912-1325831, 1471368-1472903 and  
 1492104-1493255). Subtraction by Bento Soares and M.  
 Fatima Bonaldo."

## ORIGIN

Query Match 54.8%; Score 542; DB 9; Length 572;  
 Best Local Similarity 97.7%; Pred. No. 1.6e-57;  
 Matches 559; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

```

408 CTGGGCACTTTCACGGGCTTCTCGCCCTGAGCGCTGGCGCTGGCGCGGAG
572 CTGGGCACTTTCACGGGCTTCTCGCCCTGAGCGCTGGCGCTGGCGCGGAG
468 CGGTGTGACCTTTCACGGGCTTCTCGCCCTGAGCGCTGGCGCTGGCGCGGAG
512 CGGTGTGACCTTTCACGGGCTTCTCGCCCTGAGCGCTGGCGCTGGCGCGGAG
528 CAGGCGAGCGGAGGACCAAGATCGACCTTCGGGCTGAAGCCCGCTTGAAGACCTTGGAC
452 CAGGCGAGCGGAGGACCAAGATCGACCTTCGGGCTGAAGCCCGCTTGAAGACCTTGGAC
588 GAGCTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
392 GAGCTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
648 GAGACTGTCTCCGCTTACTACAGGCGCTGTGCGAGCTGTGCGACCCGGAGGACTCTC 707

```

Db 332 GAGAACTGCTCGCTACTACAGACGCTCTGACGCTGCTGACCCCGAGGCACTCTC 273  
 QY 708 GCCGCTCTCAGAGT-CCTGTGGCGCGGAGAGTGTGCAACCTCCGAAAGGGAGCGTGC 766  
 Db 272 GCCGCTCTCAGAGTCCCTGTGGCGCGGAGAGTGTGCAACCTCCGAAAGGGAGCGTGC 213  
 QY 767 GCGCGAGTGTGTGCGAAACCTTAACGACGCAATCCGCGGAGAGTGTGCAATCAG 826  
 Db 212 GCGCGAGTGTGTGCGAAACCTTAACGACGCAATCCGCGGAGAGTGTGCAATCAG 153  
 QY 827 CTTCCGCGCGCGGAGAGTGTGCAATCCGCGGAGAGTGTGCAATCAG 886  
 Db 152 CTTCCGCGCGCGGAGAGTGTGCAATCCGCGGAGAGTGTGCAATCAG 93  
 QY 887 GAGTGGGCTCGAGGAGGAGTGTGCAATCCGCGGAGAGTGTGCAATCAG 946  
 Db 92 GAGTGGGCTCGAGGAGGAGTGTGCAATCCGCGGAGAGTGTGCAATCAG 33  
 QY 947 GAAATTAAGTGGGCTCGAGGAGGAGTGTGCAATCCGCGGAGAGTGTGCAATCAG 978  
 Db 32 GAAATTAAGTGGGCTCGAGGAGGAGTGTGCAATCCGCGGAGAGTGTGCAATCAG 1

RESULT 33  
 BG910899 639 bp mRNA linear EST 05-JUN-2001  
 LOCUS 602812540F1 NCI CGAP\_Brn67 Homo sapiens cDNA clone IMAGE:4944633  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BG910899  
 VERSION BG910899.1 GI:14291375  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 TITLE NIH-MGC http://mgi.nci.nih.gov/.  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-r@mail.nih.gov  
 Tissue Procurement: David N. Louis, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LIML)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LIML at:  
 http://image.llnl.gov  
 Plate: LIML0889 row: p column: 10  
 High quality sequence stop: 614.

FEATURES  
 source  
 1..639  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4944633"  
 /issue\_type="anaplastic oligodendroglioma with 1p/19q  
 loss"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI CGAP\_Brn67"  
 /note="Organ: Brain; Vector: pCMV-Sport6; Site: 1; Nott;  
 Site: 2; Sali; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 2.3 kb. Constructed by Life  
 Technologies. Note: this is a NCI CGAP library."

## ORIGIN

Query Match 54.6%; Score 540.2; DB 12; Length 639;  
 Best Local Similarity 97.0%; Pred. No. 2.6e-57;  
 Matches 614; Conservative 0; Mismatches 13; Indels 6; Gaps 6;  
 QY 354 GCCACGCTTGGCGAAGCGGCGGCTCATCAGCCAGAGAGCGCTGGACCTGGGC 413  
 Db 9 GCGCAGCTTGGCGAAGCGGCGGCTCATCAGCCAGAGAGCGCTGGACCTGGGC 68

QY 414 ACCTTCAGGAGCTACTCCGCGCTGAGCCTTGAGCCTTGAGC-CTGCGCGGAGAGCGGCGGT 472  
 Db 69 ACCTTCAGGAGCTACTCCGCGCTGAGCCTTGAGCCTTGAGCCTTGAGCCTTGAGCCTTGAGC 128  
 QY 473 GGTGACCTGAGAGTGAAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 532  
 Db 129 GGTGACCTGAGAGTGAAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 188  
 QY 533 CGAGGCGGAGAGCAAGATGACCTTCGCGCTGAAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 592  
 Db 189 CGAGGCGGAGAGCAAGATGACCTTCGCGCTGAAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 248  
 QY 593 GCTGCGGCGGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 652  
 Db 249 GCTGCGGCGGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 308  
 QY 653 CTGCTCGGCTACTACAGAGCGGCTGCTGAGCTGCTGAGCGGAGCGGAGCGGAGCGGAGCGG 711  
 Db 309 CTGCTCGGCTACTACAGAGCGGCTGCTGAGCTGCTGAGCGGAGCGGAGCGGAGCGGAGCGG 368  
 QY 712 TCTTCAGAGCTCTGTGCGCGGAGAGTGTGCAACCTTCGAAAGGGAGCGGAGCGGAGCGG 771  
 Db 369 TCTTCAGAGCTCTGTGCGCGGAGAGTGTGCAACCTTCGAAAGGGAGCGGAGCGGAGCGGAGCGG 428  
 QY 772 AGTGTGTGCGAAACCTTAACGACGATCCGCGGAGGAC-GTCAGGGTCTACATCAGCCTC 830  
 Db 429 AGTGTGTGCGAAACCTTAACGACGATCCGCGGAGGACGTCAGGGTCTACATCAGCCTC 488  
 QY 831 CTGCGCGGCGGAGGATGACCTACCTTGAGCTTCAAGATCTAGAGCTGCGGAGCGGAGCGG 890  
 Db 489 CTG-CCTGCGGCGGAGGATGACCTACCTTGAGCTTCAAGATCTAGAGCTGCGGAGCGGAGCGG 546  
 QY 891 GGGCTCGAGGAGGAGGATGAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 949  
 Db 547 GGGCTCGAGGAGGAGGATGAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 606  
 QY 950 AATTAAGTGGGCTGGGAGACACAAAAA 982  
 Db 607 AATTAAGTGGGCTGGGAGACACAAAAA 639

RESULT 34  
 AW163525 566 bp mRNA linear EST 09-NOV-1999  
 LOCUS AW163525  
 DEFINITION IMAGE:2784055 5' similar to TR:085769 085769 HYPOTHETICAL 24.8 KD  
 PROTEIN. ; mRNA sequence.  
 ACCESSION AW163525  
 VERSION AW163525.1 GI:6302558  
 KEYWORDS EST.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens (human)

REFERENCE  
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 TITLE Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
 Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,  
 Martin, J., Moore, B., Schellenberg, K., Stepoe, M., Tan, F.,  
 Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.  
 WashU-NCI Human EST Project  
 Unpublished (1997)  
 Other ESTs: au95h04.x1  
 Contact: Wilson RK  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: est@wustl.wustl.edu  
 This clone is available royalty-free through LIML; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seg primer: -40RP from Gibco  
 High quality sequence stop: 377.

